Lab\_2\_Worksheet

Eric Huber

02/02/2022

## Question 1:

Fdata = read.csv("FlowerData.csv")  
Fdata

## Individual Age..days. Height..cm. Colour  
## 1 a 31 5.0 purple  
## 2 b 48 16.0 yellow  
## 3 c 39 12.5 red  
## 4 d 29 6.0 red  
## 5 e 32 4.0 red  
## 6 f 37 7.0 yellow  
## 7 g 37 8.0 yellow  
## 8 h 26 5.5 yellow  
## 9 i 41 10.0 purple  
## 10 j 34 8.5 purple  
## 11 k 38 12.0 pr  
## 12 l 40 18.0 yellow  
## 13 45 16.0 yellow  
## 14 n 40 12.5 yellow  
## 15 o 43 9.5 red  
## 16 p 33 7.0 yellow  
## 17 q 35 6.0 yellow  
## 18 r 39 6.5 yellow  
## 19 s 37 12.5 red  
## 20 t 32 13.0 purple  
## 21 u 31 10.5 yellow  
## 22 v 36 11.0 red  
## 23 w 41 17.0 red  
## 24 x 39 15.5 yellow  
## 25 y 31 9.5 yellow  
## 26 z 33 10.0 yellow  
## 27 aa 33 11.0 yellow  
## 28 bb 28 5.5 red  
## 29 cc 35 13.5 red  
## 30 dd 37 16.0 yellow  
## 31 ee 42 14.0 red  
## 32 ff 45 17.0 red  
## 33 gg 37 16.5 red  
## 34 hh 46 18.0 red  
## 35 ii 37 17.5 yellow  
## 36 jj 44 19.0 yellow  
## 37 kk 44 14.0 yellow  
## 38 ll 37 8.0 yellow  
## 39 mm 29 10.0 yellow  
## 40 nn 37 9.0 red  
## 41 oo 36 12.0 purple  
## 42 pp 47 15.5 purple  
## 43 qq 38 16.0 yellow  
## 44 rr 40 16.0 red  
## 45 ss 45 17.0 purple

### (1.b)

Fdata is a data frame because it stores data tables that contains multiple data types. A matrix can only store one data type and Fdata stores numbers as well as characters therefore it must be a dataframe.

### (1.c)

FlowerMatrix = as.matrix(Fdata[ ,2:3])

### (1.d)

colnames(FlowerMatrix) <- c("Age (in days)", "Height (in cm)")

### (1.e)

rownames(FlowerMatrix) = c(Fdata$Individual)

## Question 2:

### (2.a)

mean(Fdata$Age)

## [1] 37.42222

The average age is ~37.42 days

### (2.b)

mean(Fdata$Height..cm.)

## [1] 11.87778

The average height is ~11.88 centimeters

### (2.c)

max(Fdata$Height..cm.)

## [1] 19

max(Fdata$Age..days.)

## [1] 48

The tallest flower is 19 centimeters tall and the oldest flower is 48 days old

### (2.d)

min(Fdata$Height..cm.)

## [1] 4

min(Fdata$Age..days.)

## [1] 26

The shortest flower is 4 centimeters tall and the youngest flower is 26 years old.

### (2.e)

The tallest flower from part c is yellow. The oldest flower from part c is also yellow. The shortest flower from part d is red. The youngest flower from part d is also yellow.

## Question 3:

### (3.a)

n = nrow(FlowerMatrix)

There are 45 rows in the matrix

### (3.b)

Samp = sample(1:nrow(FlowerMatrix), 15)

### (3.c)

SampleMatrix = matrix(c(FlowerMatrix[Samp, ]), nrow=15)  
colnames(SampleMatrix) <- c("Age (in days)", "Height (in cm)")

### (3.d)

colMeans(SampleMatrix)

## Age (in days) Height (in cm)   
## 36.06667 10.70000

The sample average age of the flowers is ~37.07 days and the sample average height is ~11.03 cm. These results are very similar to the results found in questions 2 a/b but they are both just a bit under the actual value. That being said depending on the sample these values could fluctuate dramatically